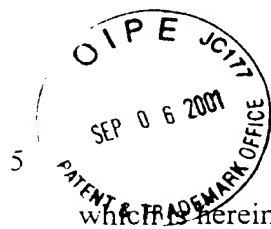


**Lithographic Apparatus, Device Manufacturing Method, and Device
Manufactured Thereby**



BACKGROUND OF THE INVENTION

5 This application claims priority from EP Application Number 00302996.4 which is herein incorporated by reference.

Field of the Invention

- 10 The present invention relates to the control of leveling, for example of the substrate and/or mask, during exposures in lithographic apparatus. More particularly, the invention relates to a system for leveling control in a lithographic projection apparatus comprising:
- a radiation system for supplying a projection beam of radiation;
 - 15 - a support structure for supporting patterning means, the patterning means serving to pattern the projection beam according to a desired pattern;
 - a substrate table for holding a substrate; and
 - a projection system for projecting the patterned beam onto a target portion of the substrate;
 - 20 - a level sensor for measuring at least one of a perpendicular position and tilt about at least one parallel axis of a surface of an object held by one of the support structure and the substrate table, and generating a position signal indicative thereof, perpendicular referring to a direction substantially perpendicular to the said surface and parallel referring to a direction substantially parallel to said surface; and
 - 25 - a servo system responsive to said position signal for moving said object to a desired position.

Description of the Related Art

- 30 The term "patterning structure", "patterning means" or "mask" as here employed should be broadly interpreted as referring to means that can be used to endow an incoming radiation beam with a patterned cross-section, corresponding to a pattern that is to be created in a target portion of the substrate; the term "light valve" can also be used in this

context. Generally, the pattern will correspond to a particular functional layer in a device being created in the target portion, such as an integrated circuit or other device (see below).

Examples of such patterning means include:

- 5 - A mask. The concept of a mask is well known in lithography, and it includes mask types such as binary, alternating phase-shift, and attenuated phase-shift, as well as various hybrid mask types. Placement of such a mask in the radiation beam causes selective transmission (in the case of a transmissive mask) or reflection (in the case of a reflective mask) of the radiation impinging on the mask, according to the pattern on the mask. In the case of a mask, the support structure will generally be a mask table, 10 which ensures that the mask can be held at a desired position in the incoming radiation beam, and that it can be moved relative to the beam if so desired.
- A programmable mirror array. An example of such a device is a matrix-addressable surface having a viscoelastic control layer and a reflective surface. The basic principle 15 behind such an apparatus is that (for example) addressed areas of the reflective surface reflect incident light as diffracted light, whereas unaddressed areas reflect incident light as undiffracted light. Using an appropriate filter, the said undiffracted light can be filtered out of the reflected beam, leaving only the diffracted light behind; in this manner, the beam becomes patterned according to the addressing pattern of the matrix-addressable surface. The required matrix addressing can be performed using 20 suitable electronic means. More information on such mirror arrays can be gleaned, for example, from United States Patents US 5,296,891 and US 5,523,193, which are incorporated herein by reference. In the case of a programmable mirror array, the said support structure may be embodied as a frame or table, for example, which may be fixed or movable as required.
- 25 - A programmable LCD array. An example of such a construction is given in United States Patent US 5,229,872, which is incorporated herein by reference. As above, the support structure in this case may be embodied as a frame or table, for example, which may be fixed or movable as required.

30 For purposes of simplicity, the rest of this text may, at certain locations, specifically direct itself to examples involving a mask and mask table; however, the general principles discussed in such instances should be seen in the broader context of the patterning means or patterning structure as hereabove set forth.

Lithographic projection apparatus can be used, for example, in the manufacture of integrated circuits (ICs). In such a case, the patterning structure may generate a circuit pattern corresponding to an individual layer of the IC, and this pattern can be imaged onto a target portion (*e.g.* comprising one or more dies) on a substrate (silicon wafer) that has been coated with a layer of radiation-sensitive material (resist). In general, a single wafer will contain a whole network of adjacent target portions that are successively irradiated via the projection system, one at a time. In current apparatus, employing patterning by a mask on a mask table, a distinction can be made between two different types of machine. In one type of lithographic projection apparatus, each target portion is irradiated by exposing the entire mask pattern onto the target portion at once; such an apparatus is commonly referred to as a wafer stepper. In an alternative apparatus — commonly referred to as a step-and-scan apparatus — each target portion is irradiated by progressively scanning the mask pattern under the projection beam in a given reference direction (the "scanning" direction) while synchronously scanning the substrate table parallel or anti-parallel to this direction; since, in general, the projection system will have a magnification factor M (generally < 1), the speed V at which the substrate table is scanned will be a factor M times that at which the mask table is scanned. More information with regard to lithographic devices as here described can be gleaned, for example, from US 6,046,792, incorporated herein by reference.

In a manufacturing process using a lithographic projection apparatus, a pattern (*e.g.* in a mask) is imaged onto a substrate that is at least partially covered by a layer of radiation-sensitive material (resist). Prior to this imaging step, the substrate may undergo various procedures, such as priming, resist coating and a soft bake. After exposure, the substrate may be subjected to other procedures, such as a post-exposure bake (PEB), development, a hard bake and measurement/inspection of the imaged features. This array of procedures is used as a basis to pattern an individual layer of a device, *e.g.* an IC. Such a patterned layer may then undergo various processes such as etching, ion-implantation (doping), metallization, oxidation, chemo-mechanical polishing, etc., all intended to finish off an individual layer. If several layers are required, then the whole procedure, or a variant thereof, will have to be repeated for each new layer. Eventually, an array of devices will be present on the substrate (wafer). These devices are then separated from one another by a technique such as dicing or sawing, whence the individual devices can be mounted on a carrier, connected to pins, etc. Further information regarding such processes can be obtained, for example, from the book "Microchip Fabrication: A Practical Guide to Semiconductor

Processing", Third Edition, by Peter van Zant, McGraw Hill Publishing Co., 1997, ISBN 0-07-067250-4, incorporated herein by reference.

For the sake of simplicity, the projection system may hereinafter be referred to as the "lens"; however, this term should be broadly interpreted as encompassing various types of projection system, including refractive optics, reflective optics, and catadioptric systems, for example. The radiation system may also include components operating according to any of these design types for directing, shaping or controlling the projection beam of radiation, and such components may also be referred to below, collectively or singularly, as a "lens". Further, the lithographic apparatus may be of a type having two or more substrate tables (and/or two or more mask tables). In such "multiple stage" devices the additional tables may be used in parallel, or preparatory steps may be carried out on one or more tables while one or more other tables are being used for exposures. Twin stage lithographic apparatus are described, for example, in US 5,969,441 and WO 98/40791, incorporated herein by reference.

Until very recently, lithographic apparatus contained a single mask table and a single substrate table. However, machines are now becoming available in which there are at least two independently moveable substrate tables; see, for example, the multi-stage apparatus described in International Patent Applications WO98/28665 and WO98/40791. The basic operating principle behind such multi-stage apparatus is that, while a first substrate table is at the exposure position underneath the projection system for exposure of a first substrate located on that table, a second substrate table can run to a loading position, discharge a previously exposed substrate, pick up a new substrate, perform some initial measurements on the new substrate and then stand ready to transfer the new substrate to the exposure position underneath the projection system as soon as exposure of the first substrate is completed; the cycle then repeats. In this manner it is possible to increase substantially the machine throughput, which in turn improves the cost of ownership of the machine. It should be understood that the same principle could be used with just one substrate table which is moved between exposure and measurement positions.

During exposure processes, it is important to ensure that the mask image is correctly focussed on the substrate. Conventionally this has been done by measuring the vertical position of the best focal plane of the aerial image of the mask pattern relative to the projection lens before an exposure or a series of exposures. During each exposure, the vertical position of the upper surface of the substrate relative to the projection lens is measured and

the position of the substrate table is adjusted so that the substrate surface lies in the best focal plane. However, known leveling systems have not always been able to effect sufficiently accurate positioning of the substrate surface in the best focal plane and can cause undesirable X and Y movements of the substrate due to cross-talk from Rx and Ry leveling adjustments.

- 5 Such X and Y movements cause overlay errors which are particularly undesirable.

SUMMARY OF THE INVENTION

An object of the present invention is to provide a control system capable of improved "on-the-fly" leveling (that is leveling based on position measurements made during the exposure rather than in advance) performed on a substrate or mask in a lithographic projection apparatus during exposure processes and in particular to reduce focussing errors, cross-talk between tilts and horizontal shifts and unnecessary object table movements.

This and other objects are achieved according to the invention in a lithographic projection apparatus as specified in the opening paragraph, characterized by a filter connected between said level sensor and said servo system for filtering said position signal.

The present invention, by interposing a filter between the level sensor and the servo system for leveling, enables improvements in the leveling performance. In particular, undesirable movements to follow high spatial frequency (height) variations in the substrate surface can be avoided. Also, trade-offs between performance in different degrees of freedom can be made, especially to avoid cross-talk into horizontal displacements of the substrate which would result in overlay errors. Preferably, the level sensor, optionally in cooperation with a position sensor such as an interferometer or a Linear Variable Differential Transformer (LVDT) measurement system, generates a setpoint which the servo system aims to follow.

25 The filter then filters that setpoint.

According to a further aspect of the present invention there is provided a device manufacturing method comprising the steps of:

- providing a substrate that is at least partially covered by a layer of radiation-sensitive material;
- 30 providing a projection beam of radiation using a radiation system;
- using patterning means to endow the projection beam with a pattern in its cross-section;
- measuring at least one of a perpendicular position and tilt about at least one

parallel axis of a surface of an object held by one of said support structure and said substrate table, and for generating a position signal indicative thereof, perpendicular referring to a direction substantially perpendicular to the said surface and parallel referring to a direction substantially parallel to said surface;

5 providing a servo system responsive to said position signal for moving said object to a desired position; and

 projecting the patterned beam of radiation onto a target portion of the layer of radiation-sensitive material whilst operating said servo system to maintain said object at said desired position;

10 characterized by the step of:

 filtering said position signal before it is used by said servo system to control the position of said object.

 Although specific reference may be made in this text to the use of the apparatus according to the invention in the manufacture of ICs, it should be explicitly
15 understood that such an apparatus has many other possible applications. For example, it may be employed in the manufacture of integrated optical systems, guidance and detection patterns for magnetic domain memories, liquid-crystal display panels, thin-film magnetic heads, etc. The skilled artisan will appreciate that, in the context of such alternative applications, any use of the terms "reticle", "wafer" or "die" in this text should be considered
20 as being replaced by the more general terms "mask", "substrate" and "target portion ", respectively.

 In the present document, the terms "radiation" and "beam" are used to encompass all types of electromagnetic radiation, including ultraviolet (UV) radiation (*e.g.* with a wavelength of 365, 248, 193, 157 or 126 nm) and extreme ultra-violet (EUV or XUV)
25 radiation, *e.g.* having a wavelength in the range 5-20 nm), as well as particle beams, such as ion beams or electron beams.

BRIEF DESCRIPTION OF THE DRAWINGS

30 The present invention will be described below with reference to exemplary embodiments and the accompanying schematic drawings, in which:

 Figure 1 depicts a lithographic projection apparatus according to a first embodiment of the invention;

Figure 2 depicts a level sensor device used in the first embodiment of the invention;

Figure 3 is a diagram of a control system used in the first embodiment of the invention;

5 Figure 4 is a diagram used to explain measurements used in a second embodiment of the invention;

Figure 5 is a diagram of a control system used in a second embodiment of the invention;

10 Figure 6 is a diagram showing the relative positions of measurement spots used in a third embodiment of the invention;

Figure 7 is a diagram of a control system used in the third embodiment of the invention;

Figure 8 is a diagram of a control system used in a fourth embodiment of the invention;

15 Figure 9 is a diagram of a control system used in a fifth embodiment of the invention;

Figure 10 is a diagram of a control system used in a sixth embodiment of the invention;

20 Figures 11A and B are graphs showing Z positions of a wafer table during scanning of a test wafer with a conventional apparatus and with an apparatus embodying the invention;

Figures 12A and B are graphs showing Ry positions of a wafer table during scanning of a test wafer with a conventional apparatus and with an apparatus embodying the invention;

25 Figures 13A and B are graphs showing Z position level sensor transfer functions of a conventional level sensor and a level sensor with filtering according to the invention, in each case compared to an ideal level sensor;

Figures 14A and B are graphs showing Ry position level sensor transfer functions of a conventional level sensor and a level sensor with filtering according to the invention, in each case compared to an ideal level sensor; and

30

Figure 15 is a table of wafer shape filter settings in two examples of the invention.

In the drawings, like references indicate like parts.

DETAILED DESCRIPTION OF THE INVENTION

Embodiment 1

Figure 1 schematically depicts a lithographic projection apparatus according to a particular embodiment of the invention. The apparatus comprises:

- a radiation system Ex, IL, for supplying a projection beam PB of radiation (*e.g.* UV or EUV radiation, electrons or ions). In this particular case, the radiation system also comprises a radiation source LA;
- a first object table (mask table) MT provided with a mask holder for holding a mask MA (*e.g.* a reticle), and connected to first positioning means for accurately positioning the mask with respect to item PL;
- a second object table (substrate table) WT provided with a substrate holder for holding a substrate W (*e.g.* a resist-coated silicon wafer), and connected to second positioning means for accurately positioning the substrate with respect to item PL;
- a projection system ("lens") PL (*e.g.* a refractive or catadioptric system, a mirror group or an array of deflectors) for imaging an irradiated portion of the mask MA onto a target portion C (*e.g.* comprising one or more dies) of the substrate W.

As here depicted, the apparatus is of a transmissive type (*i.e.* has a transmissive mask). However, in general, it may also be of a reflective type, for example (with a reflective mask). Alternatively, the apparatus may employ another kind of patterning means, such as a programmable mirror array of a type as referred to above.

The source LA (*e.g.* a Hg lamp, excimer laser, an undulator provided around the path of an electron beam or storage ring or synchrotron, a laser-produced plasma source, a discharge source or an electron or ion beam source) produces a beam of radiation. This beam is fed into an illumination system (illuminator) IL, either directly or after having traversed conditioning means, such as a beam expander Ex, for example. The illuminator IL may comprise adjusting means AM for setting the outer and/or inner radial extent (commonly referred to as σ -outer and σ -inner, respectively) of the intensity distribution in the beam. In addition, it will generally comprise various other components, such as an integrator IN and a condenser CO. In this way, the beam PB impinging on the mask MA has a desired uniformity and intensity distribution in its cross-section.

It should be noted with regard to Figure 1 that the source LA may be within the housing of the lithographic projection apparatus (as is often the case when the source LA

is a mercury lamp, for example), but that it may also be remote from the lithographic projection apparatus, the radiation beam which it produces being led into the apparatus (e.g. with the aid of suitable directing mirrors); this latter scenario is often the case when the source LA is an excimer laser. The current invention and Claims encompass both of these scenarios.

The beam PB subsequently intercepts the mask MA, which is held on a mask table MT. Having traversed the mask MA, the beam PB passes through the lens PL, which focuses the beam PB onto a target portion C of the substrate W. With the aid of the second positioning means (and interferometric measuring means IF), the substrate table WT can be moved accurately, *e.g.* so as to position different target portions C in the path of the beam PB. Similarly, the first positioning means can be used to accurately position the mask MA with respect to the path of the beam PB, *e.g.* after mechanical retrieval of the mask MA from a mask library, or during a scan. In general, movement of the object tables MT, WT will be realized with the aid of a long-stroke module (course positioning) and a short-stroke module (fine positioning), which are not explicitly depicted in Figure 1. However, in the case of a wafer stepper (as opposed to a step-and-scan apparatus) the mask table MT may just be connected to a short-stroke actuator, or may be fixed.

The depicted apparatus can be used in two different modes:

1. In step mode, the mask table MT is kept essentially stationary, and an entire mask image is projected at once (*i.e.* a single "flash") onto a target portion C. The substrate table WT is then shifted in the x and/or y directions so that a different target portion C can be irradiated by the beam PB;
2. In scan mode, essentially the same scenario applies, except that a given target portion C is not exposed in a single "flash". Instead, the mask table MT is movable in a given direction (the so-called "scan direction", *e.g.* the y direction) with a speed v , so that the projection beam PB is caused to scan over a mask image; concurrently, the substrate table WT is simultaneously moved in the same or opposite direction at a speed $V = Mv$, in which M is the magnification of the lens PL (typically, $M = 1/4$ or $1/5$). In this manner, a relatively large target portion C can be exposed, without having to compromise on resolution.

An important factor influencing the imaging quality of a lithographic apparatus is the accuracy with which the mask pattern image is focused on the substrate. In practice, since the scope for adjusting the position of the focal plane of the projection system PL is limited and the depth of focus of that system is small, this means that the target portion of the

wafer (substrate) must be positioned precisely in the focal plane of the projection system PL. To do this, it is of course necessary to know both the position of the focal plane of the projection system PL and the position of the top surface of the wafer. Wafers are generally polished to a very high degree of flatness but nevertheless deviation of the wafer surface from perfect flatness (referred to as "unflatness") of sufficient magnitude to affect focus accuracy can occur. Unflatness may be caused, for example, by variations in wafer thickness, distortion of the shape of the wafer or contaminants on the wafer table. The presence of structures due to previous process steps also significantly affects the wafer height (flatness). In the present invention, the cause of unflatness is largely irrelevant; only the height of the top surface of the wafer is considered. Unless the context otherwise requires, references below to "the wafer surface" refer to the top surface of the wafer onto which will be projected the mask image.

During exposures, the position and orientation of the wafer surface relative to the projection optics PL are measured. The perpendicular, or vertical, position (Z) and parallel, or horizontal, tilts (R_x , R_y) of the wafer table WT are adjusted to keep the wafer surface at the optimal focus position. The perpendicular, or vertical, position refers to the position along an axis substantially perpendicular to the plane of the wafer surface, and the parallel, or horizontal, tilts refer to tilts along at least one axis parallel to the plane of the wafer surface. The detector, referred to herein as the level sensor, which may be used for this is shown in Figure 2. It comprises a radiation source S which has two emitting areas S1, S2 and supplies two beams, a reference beam and a measurement beam having a wide wavelength band. Also shown are an object grating G1 and an image grating G2. Optical systems (depicted for clarity as simple lenses) L1 and L2 image the object grating G1 onto the image grating G2, the reference beam having been reflected by the outer surface RP of the projection optics PL and the measurement beam by the wafer surface. Detectors DE2, DE1 behind the image grating G2 give, when irradiated, signals which can be measured by a meter ME or other suitable instrument, indicative of the relative positions of the points where the reference beam and measurement beam are reflected by the projection optics PL and wafer surface respectively. By using more than one such system, e.g. four, the relative heights of a corresponding number of points on the wafer surface can be measured and local tilts of the wafer surface determined. The level sensor is described in greater detail in EP-0 502 583-A and US 5,191,200, for example, which documents are incorporated herein by reference.

A schematic of the leveling control system is shown in Figure 3. The physical

components of the system are the level sensor LS, wafer shape filter WSF and servo system SV. The servo system SV is a closed-loop system, including necessary control circuitry, a mechanism for driving the wafer table and a positioning system. The level sensor output ls is filtered by the wafer shape filter WSF to give a filtered signal ls' which forms the setpoint of the servo system. The servo system drives the wafer table WT to a vertical position vp and may introduce a horizontal servo error hse in the horizontal position of the wafer. Such an error can be caused by a non-zero Abbe arm for the Rx and Ry rotations carried out by the servo system, for example. In other words, the servo system rotates the wafer table about axes not lying exactly on the wafer surface. Any error vse in the vertical position signals output by the servo system SV can be measured by subtracting the filtered level sensor signal ls' from the measured vertical position vp (vp is measured using an interferometer or LVDT, for example). These components and their interconnections are shown in solid in Figure 3. It should be noted that leveling of the wafer is carried out in three degrees of freedom; vertical (Z) position and rotation about orthogonal horizontal axes (Rx and Ry). Figure 3, and some later Figures showing other embodiments of the invention, show general control architectures applicable for all three degrees of freedom Z, Rx and Ry. Unless the context otherwise requires, signals such as ls , ls' , Zif , etc., include data of those three degrees of freedom.

The transfer function H_{ls} of the level sensor LS is not ideal. If an ideal level sensor ILS is notionally introduced into the system then the various possible errors in the complete system can be identified. Since the ideal level sensor cannot be built, such an ideal sensor and the errors derived by reference to it are shown in phantom in Figure 3. These errors are the level sensor error lse , dynamic measurement error dme and dynamic leveling error dle . Thus the errors in the leveling system may be defined as:

$$lse = \text{wafer} \cdot H_{ils} - \text{wafer} \cdot H_{ls} \quad (1)$$

$$dme = \text{wafer} \cdot H_{ils} - \text{wafer} \cdot H_{ls} \cdot H_{wsf} \quad (2)$$

$$vse = \text{wafer} \cdot H_{ls} \cdot H_{wsf} - \text{wafer} \cdot H_{ls} \cdot H_{wsf} \cdot H_{sv} \quad (3)$$

$$dle = \text{wafer} \cdot H_{ils} - \text{wafer} \cdot H_{ls} \cdot H_{wsf} \cdot H_{sv} \quad (4)$$

$$hse = \text{wafer} \cdot H_{ls} \cdot H_{wsf} \cdot H_{sv} \quad (5)$$

where H_{aa} is the transfer function of element AA in the control system.

These various transfer functions will in general be functions of Z, Rx and Ry and may include terms representing cross-talk into other degrees of freedom. Of these errors, the first four are defined for Z, Rx, Ry and Ztotal, the last only for X and Y. Ztotal is a combination of Z, Rx and Ry errors in such a way that it represents the maximal Z displacement in the radiation

system's exposure slit, effectively the maximum Z error on one of the four corners of the slit. Z_{total} is calculated as $Z \pm R_x \cdot \text{slitsize}_Y / 2 \pm R_y \cdot \text{slitsize}_X / 2$. Slitsize_Y is defined as the width of said projection beam in a scanning direction of one of the support structure and the wafer table, and slitsize_X refers to a width of said projection beam in a direction substantially perpendicular to the said scanning direction.

The transfer function H_{wsf} of the wafer shape filter is determined for each application to provide the desired improvements to the above errors. For example, the transfer function H_{wsf} may be empirically derived to compensate for the divergence of the transfer function H_{ls} of the actual level sensor LS and so reduces the dynamic measurement error dme to zero. The ideal level sensor transfer function has a magnitude that decreases with spatial frequency, and a first zero-crossing at a spatial frequency equal to the inverse of the width of the exposure slit in the scan direction (in the case of a step-and-scan apparatus). This is advantageous as it prevents the wafer table attempting to follow variations in the wafer surface of wavelength shorter than the slit width and in particular reduces undesired horizontal movements due to high-frequency cross-talk.

The wafer shape filter transfer function can also be adjusted to compensate for or compromise between the other errors. Appropriate forms for the wafer shape filter transfer function to achieve the desired effects can be derived empirically or by modeling the servo system. For example, in one servo system it was determined that Y errors were out of specification whilst Z_{total} and Rx errors were comfortably within limits. A notch filter in the Rx wafer shape filter transfer function with a center frequency equal to the peak frequency of the Y moving average error was found to improve Y accuracy at an acceptable cost to Rx and Z_{total} . The damping coefficients were selected to provide the desired improvement in Y while reducing the cost to Rx and Z_{total} .

Embodiment 2

In a second embodiment, shown in Figure 5, the control system makes use of information indicating the position of the wafer table WT provided by an interferometric displacement measurement system IF. Suitable three, five and six-axis interferometric metrology systems are described in WO99/28790 and WO99/32940, for example, which are incorporated herein by reference. An LVDT measurement system may also be used in place of the interferometer. In such a system, three LVDTs are located under the wafer table WT

and their outputs transformed to give Z , R_x and R_y data. As shown in Figure 4, the interferometer system IF measures the position Z_{if} of the wafer table WT (sometimes referred to as the mirror block, as the interferometer system makes use of mirrors bonded to the sides of the wafer table) relative to the focal plane FP of the projection lens system PL while the level sensor measures the height l_s of the upper surface of the wafer W. (Note that while the l_s and Z_{if} measurements are shown spaced apart in Figure 4 for clarity, in fact the interferometer and level sensor should be arranged to make measurements at the same position in the XY plane.) The interferometer data, though denoted simply Z_{if} , includes information regarding the horizontal tilt, R_x and R_y , of the wafer table as well as vertical position, Z . By subtracting the level sensor data from the interferometer data, a value for the wafer shape ws is obtained, i.e.:

$$ws = Z_{if} - l_s \quad (6)$$

The control system using the interferometer data is shown in Figure 5. The control strategy for this system is that the wafer shape filter WSF provides the filtered wafer shape signal ws' which acts as setpoint data for an inner closed-loop control system (within the double dotted line in Figure 5) comprising controller CONT, the short-stroke table drive system MECH, the interferometer IF and a subtractor which subtracts the position of the wafer table as indicated by the interferometer data Z_{if} from the filtered wafer shape data ws' . In the second embodiment, the wafer shape filter WSF acts on the wafer shape data ws (which represents the actual shape of the wafer) rather than the level sensor data (which includes the instantaneous position of the wafer table). The inner loop has a high bandwidth, e.g. 50 or 100 Hz or more, and is able to follow the wafer shape setpoint ws' accurately. The outer loop determines the setpoint by filtering the wafer shape signal ws . The wafer shape filter WSF will therefore not affect the performance of the inner loop. The outer loop needs to be stable and to have limited closed-loop amplification.

As in the first embodiment, the wafer shape filter is selected to correct measurement errors in the level sensor LS and to reduce vertical (tilt) to horizontal cross-over.

30 Embodiment 3

A third embodiment of the invention is described with reference to Figures 6 and 7. The third embodiment incorporates so-called "look-ahead" in the level sensor LS to

compensate for delay which is caused in the wafer shape filter. The level sensor including look-ahead is denoted LS' and utilizes a measurement spot pattern as depicted in Figure 6. Measurement spots P1 and Q2 are positioned ahead of the center of the projection lens while Q1 and P2 are behind. The corresponding signals are denoted as ZP1, ZQ2, ZQ1, ZP2, respectively. With this 4-spot layout, sensor look-ahead for Z and Ry position is effected by weighting the advance spot measurements more heavily than the back spots. (Note that sensor look-ahead is not used in Rx because Rx measurements require both advance and back spot measurements.) Without sensor look-ahead, center level sensor Z, Rx and Ry signals are calculated as follows:

$$ls_centZ = (ZP1 + ZP2 + ZQ1 + ZQ2)/4 \quad (7)$$

$$ls_centRx = ((ZP1 + ZQ2)/2 - (ZP2 + ZQ1)/2)/arm_y \quad (8)$$

$$ls_centRy = ((ZP1 + ZQ1)/2 - (ZP2 + ZQ2)/2)/arm_x \quad (9)$$

To calculate look-ahead in Z and Ry, gradient values are defined as follows:

$$ls_gradZ = dz/dy = ls_centRx \quad (10)$$

$$ls_gradRx = 0 \quad (11)$$

$$ls_gradRy = dRy/dy = ((ZP1 - ZQ2)/arm_x - (ZQ1 - ZP2)/arm_x)/arm_y \quad (12)$$

The look-ahead level sensor readings are then:

$$ls_frontZ = ls_centZ + y_l_aZ \cdot ls_gradZ \quad (13)$$

$$ls_frontRx = ls_centRx \quad (14)$$

$$ls_frontRy = ls_centRy + y_l_aRy \cdot ls_gradRy \quad (15)$$

where y_l_a is the look-ahead distance which can be different for Z and Ry.

The control system, shown in Figure 7, is then essentially the same as that of the second embodiment, shown in Figure 5, save that the level sensor LS' is adapted to provide the gradient signals, and a look-ahead multiplier y_l_a and adder are introduced to generate the sensor look-ahead data.

30 Embodiment 4

The fourth embodiment, which is shown in Figure 8, is similar to the third but includes look-ahead in the interferometer IF, or LVDT measurement system, as well. This

avoids errors in ls_frontZ which may occur in the third embodiment when there is a significant Rx tilt. In the third embodiment, the Z level sensor front signal and the Z interferometer signal are not measured at exactly the same spot so that there will be an error in the Z wafer shape signal ws if there is a significant Rx tilt. Accordingly, an interferometer
 5 gradient is defined for the Z signal, as follows:

$$ifm_gradZ = ifm_centRx \quad (16)$$

The forward-measured interferometer Z signal is then:

$$ifm_frontZ = ifm_centZ + y_l_aZ \cdot ifm_gradZ \quad (17)$$

10

Note that the interferometer gradients for Rx and Ry are defined as zero so that the corresponding look-ahead signals are equal to the center signals.

The resulting control system architecture is shown in Figure 8; it corresponds to that of the third embodiment save for the additional multiplier and adder to generate the
 15 if_front signals.

Embodiment 5

The control system architecture of a fifth embodiment of the invention is
 20 shown in Figure 9. This arrangement is effectively the same as the fourth embodiment but, by subtracting the center and gradient signals before multiplication by y_l_a , one multiplier is saved.

Embodiment 6

25

A sixth embodiment of the invention is shown in Figure 10. The sixth embodiment incorporates an additional correction AF_corr to compensate for changes in the position of the actual best focal plane. Such changes may be effected deliberately or may be caused by temperature variations in the elements of the projection optical system PL and
 30 temperature or pressure variation in the gas or air filling the projection optical system PL. A measured or predicted change of the actual focal plane in Z or Ry is automatically compensated for in the level sensor LS' which measures the position of the wafer surface relative to the optimum focal plane. However, a change of the optimum focal plane in Rx

will, with sensor look-ahead, result in an error in the Z position of the wafer surface. To prevent this, the wafer shape Z value is corrected by $-\Delta Rx \cdot y_{l_a}$, where ΔRx is the change in the position of the optimum focal plane, or the Z gradient is corrected by $-\Delta Rx$. The latter alternative is effected in the sixth embodiment in which AF_corr is subtracted from the differential gradient signal $if_grad - ls_grad$. AF_corr is defined as ΔRx for Z and zero for Rx and Ry.

Examples

To demonstrate the effectiveness of the present invention the servo architecture of the sixth embodiment was used with a fourth order wafer shape filter comprising two second order notch filters. The filter and look-ahead settings for Z, Rx and Ry for two examples, Example 1 and Example 2, are shown in Figure 15. In Figure 15, "nu" indicates "not used" and "na" indicates "not available".

In Example 1, no Rx wafer shape filter was used and the wafer shape filter acts on a time series of values representing heights at positions spaced in the Y (scanning) direction. However in Example 2, an Rx filter was added to the filter of example 1 to improve the Y performance at the expense of Rx performance. Simulations were carried out using test data derived from a sample of six test wafers. In the simulations, moving averages (MA) and moving standard deviations (MSD) for servo errors in Z_{total} , Z, Rz, Ry, X and Y, as well as dynamic leveling errors in Z_{total} , Z, Rx and Ry, were calculated, i.e. a total of 120 values. As compared to leveling without any wafer shape filtering, Example 1 reduced the number of out-of-spec results from 20 to 11 whilst Example 2 reduced this to 1.

The wafer shape filter settings of Examples 1 and 2 are based on a scanning speed of 250mm/s. For other scanning speeds, the look-ahead distances and filters can be adapted, e.g. so as to maintain a constant look-ahead time, rather than distance. Similarly, the frequency values of the wafer shape filter can be made proportional to scanning speed so that they represent constant spatial frequencies.

The effectiveness of the present invention is further demonstrated by Figures 11 to 14 which show test results obtained using the filter of Example 1 and a test wafer with a special (waved) step topology. In the negative X half of the wafer the surface has a step topology with decreasing wavelength in the Y direction. The positive X half is flat. Figures 11A and 11B show actual Z position movements (dashed) on this wafer compared to ideal (solid), without and with wafer shape filtering respectively. Figures 12A and 12B show actual

Ry movements (dashed) compared to ideal (solid), without and with wafer shape filtering respectively. Figures 13A and 13B show actual Z level sensor transfer functions (dashed) compared to ideal (solid), without and with wafer shape filtering respectively. Figures 14A and 14B show actual Ry level sensor transfer functions (dashed) compared to ideal (solid), without and with wafer shape filtering respectively. It can readily be seen that with the invention the transfer functions and wafer movements are considerably closer to the ideal. In particular, undesirable high-frequency movements of the wafer table are avoided.

As mentioned above, the actual form of the filter will be determined according to the specific embodiment of the invention and the desired performance criteria. One approach to selection of a suitable filter is to first find a level sensor look-ahead distance which ensures that the look-ahead transfer function of the level sensor lies above the ideal transfer function, at least up to the first zero-crossing at $1/\text{slitsize}Y$. Using a two-notch filter, the first notch is then used to shape the transfer function up to the first zero-crossing. The second notch is used to filter off the frequencies higher than the first zero-crossing and to adjust the phase of the transfer function up to the first zero-crossing.

While we have described above specific embodiments of the invention it will be appreciated that the invention may be practiced otherwise than as described. The description is not intended to limit the invention. It should be explicitly noted that the current invention can be applied to substrate leveling alone, to mask leveling alone, or to a combination of substrate leveling and mask leveling.